

# Northumbria Research Link

Citation: McClory, James, Timson, David J., Singh, Warispreet, Zhang, Jian and Huang, Meilan (2017) Reaction Mechanism of Isopentenyl Phosphate Kinase: A QM/MM Study. The Journal of Physical Chemistry B, 121 (49). pp. 11062-11071. ISSN 1520-6106

Published by: American Chemical Society

URL: <https://doi.org/10.1021/acs.jpcb.7b08770>  
<<https://doi.org/10.1021/acs.jpcb.7b08770>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/43064/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

# **The Reaction Mechanism of Isopentenyl Phosphate Kinase: A QM/MM Study**

James McClory,<sup>†</sup> David J. Timson,<sup>‡</sup> Warispreet Singh<sup>†, §</sup>, Jian Zhang,<sup>//</sup> Meilan Huang<sup>†,\*</sup>

<sup>†</sup> School of Chemistry and Chemical Engineering, Queen's University Belfast, David Keir Building, Stranmillis Road, Belfast, BT9 5AG, Northern Ireland, United Kingdom

<sup>‡</sup> School of Pharmacy and Biomolecular Sciences, The University of Brighton, Huxley Building, Lewes Road, Brighton, BN2 4GJ, United Kingdom

<sup>§</sup> Almac Sciences, Almac House, 20 Seagoe Industrial Estate, Craigavon BT63 5QD, Northern Ireland, United Kingdom

<sup>//</sup> 280 Chongqing Road, Shanghai Jiaotong University, Shanghai 200025, China

## ABSTRACT

Isopentenyl phosphate kinase (IPK) catalyzes the  $\text{Mg}^{2+}$ -ATP dependent phosphorylation reactions to produce isopentenyl diphosphate, an important precursor in the synthesis of isopentenols. However, the position of the divalent metal ion in the crystal structures of IPK in complex with ATP and its native substrate IP has not been definitively resolved, and as a result ambiguity surrounds the catalytic mechanism of IP, limiting its exploitation as a biofuel and in drug design. Here we report the catalytically competent structure in complex with the metal ion  $\text{Mg}^{2+}$  and elucidate the phosphorylation reaction mechanism using molecular dynamic simulations and density functional theory-based quantum mechanics/molecular mechanics calculations (B97d/AMBER99). Comparing the substrate-bound and substrate-free IPK complexes, we observed that substrate binding results in significant conformational change of three residues Lys204, Glu207, Lys211 located on the  $\alpha$ G helix to form a strong salt bridge network with Asp145, which in turn tethers the invariant Ser142 via H-bond interaction. The conformational change shuts the substrate entrance channel formed between the  $\alpha$ G and  $\alpha$ E helices. Further, we demonstrate the phosphorylation reaction occurs with a reaction barrier of 17.58 kcal/mol, which is in agreement with the previous experimental kinetic data. We found that a highly conserved Gly8 on a glycine-rich loop, together with Lys14 stabilize the transition state.

## 1. INTRODUCTION

Isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) are important precursors of isopentenol, an attractive alternative to non-renewable fuels.<sup>1</sup> The biosynthesis of isopentenols provides a promising approach for the production of these environmentally benign biofuels. Furthermore, IPP is an imperative intermediate in the biosynthesis of isoprenoids, the largest group of natural products comprised of more than 25,000 compounds. Gram-positive bacteria require IPP as a source of these isoprenoids for essential cell development, and therefore the enzymes of the mevalonate pathway have emerged as important targets for the development of antimicrobial agents when antibacterial resistance is becoming increasingly profound.<sup>2</sup>

Biosynthesis of IPP can be achieved from Mevalonate (MVA) pathway via  $Mg^{2+}$ -ATP dependent phosphorylation reactions. So far, the crystal structures of five kinases in the MVA pathway have been reported including mevalonate kinase (MVK, EC 2.7.1.36), phosphomevalonate kinase (PMK), diphosphomevalonate decarboxylase (MDD, EC 4.1.1.33) mevalonate-3-kinase (M3K, EC 2.7.1.186) and isopentenyl phosphate kinase (IPK; EC 2.7.4.26). MVK, PMK and MDD catalyze the phosphorylation of their respective substrates in the classical MVA pathway while M3K and IPK are involved in an alternative metabolic pathway to yield IPP. IPK is a member of the amino acid kinase (AAK) superfamily, characterised by a three layered protein architecture commonly referred to as the  $\alpha\beta\alpha$  sandwich, whereas the other four kinases belong to GHMP family characteristic with a phosphate binding loop (P-loop).

The AAK family is classified into two different subdivisions, the 'carboxylate' subdivision composed of N-Acetylglutamate Kinase (NAGK), carbamate kinase (CK), glutamate 5-kinase (G5K) and aspartokinase (AK) which catalyses the phosphorylation of carboxylate or

carbamate functional groups and the ‘phosphate’ subdivision comprising uridylate kinase (UK), fosfomycin resistance kinase (FomA) and IPK, which catalyses the phosphoryl transfer to a phosphate or phosphonate functional group.<sup>3</sup>

Although the crystal structures of IPK have been resolved in complex with its native substrate and cofactor ATP, the location of  $Mg^{2+}$  is either undetermined or elusive.<sup>4</sup> In order to elucidate the catalytic mechanism of IPK it is essential to locate the position of magnesium due to the critical effect of metal ion on the reaction catalyzed by the kinase.

Previously the catalytic mechanism of PMK, a GHMP kinase in the mevalonate pathway, was revealed using a combined QM/MM study.<sup>5</sup> It was found that reaction occurred via a direct phosphorylation mechanism and a positively charged residue, Lys101, stabilizes the negative charge at the cleaved  $\beta$ -, $\gamma$ -bridging oxygen of ATP. In the present work, we report determine the position of  $Mg^{2+}$  and elucidate the reaction mechanism of IPK using a molecular dynamics (MD) study in conjunction with QM/MM. Such crucial information on the reaction mechanism and key interactions in IPK provides the basis for bioengineering of the kinase for applications such as broadening the specificity of the enzyme or for the enhanced production of biofuels.

## 2. MATERIALS AND METHODS

### Protein Preparation

The crystal structure of *Thermoplasma acidophilum* (THA) IPK in complex with ATP and isopentenyl phosphate was used as the initial model (PDB code: 3LKK).<sup>4</sup> Only the chain A was kept. Magnesium was manually added to the protein near the  $\alpha/\beta/\gamma$  phosphates of ATP by superimposing with a structurally similar homologue FomA (PDB Code: 3QUN).<sup>6</sup> The missing segment 189-201 was built using Modeller v9.17.<sup>7</sup> The protonation states of

titratable residues were assigned based on analysis from the proPka server <sup>8,9</sup> and visual inspection of the local protein environment. The neutral form of His50 was used.

## **Molecular Dynamics Simulations**

The AMBER 14 (University of California, USA) package was used for Molecular Dynamic Simulations. <sup>10</sup> ATP and isopentenyl phosphate were optimized with the Gaussian 09 package <sup>11</sup> (Gaussian, Inc.) using HF/6-31g(d). RESP charges were subsequently obtained using Antechamber, a tool encoded in AMBER14 (Table S1).

The protein model was soaked with the TIP3P water box with a minimum distance of 10 Å to the boundary of the protein. The size of the box was 76.409 X 69.334 X 77.685 Å<sup>3</sup> constituting a total of 9,555 water molecules. Six sodium ions were added to neutralize the system. An initial minimization restraining the protein, ATP and substrate IP with a force constant of 50 kcal mol<sup>-1</sup>Å<sup>-2</sup> was performed to relieve any van der Waals forces between the solvent water molecules and the protein. This was followed by a 2,500 steps of unconstrained minimization of the entire system whereby 1,250 steps of the steps descent method was followed by 1,250 steps of the conjugate gradient minimization.

The system was then heated slowly from 0 to 300 K for 250 ps using the Langevin dynamics with a collision frequency of 1.0 ps<sup>-1</sup>. An equilibration simulation with no harmonic restraints applied was ran with a *NVT* ensemble and a periodic boundary condition for a further 50 ps at 300 K. A cut-off distance of 8 Å was set for non-bonded Van der Waals force while the Particle Mesh Ewald method was used to account for the long-range electrostatic interactions. The SHAKE method was utilized to fix the covalent bonds associated with the hydrogen atoms within the system. <sup>12</sup> Four replicas of 100 ns production simulation for the kinase complex with a time step of 2 fs was performed utilizing the *NPT* ensemble at 300 K and a pressure of 1 atm. RMSD was performed on Cα carbons of the protein backbone.

The mutants of IPK were generated using Discovery Studio (version 2.5) and MD simulations were run for 100 ns for each of the variant.

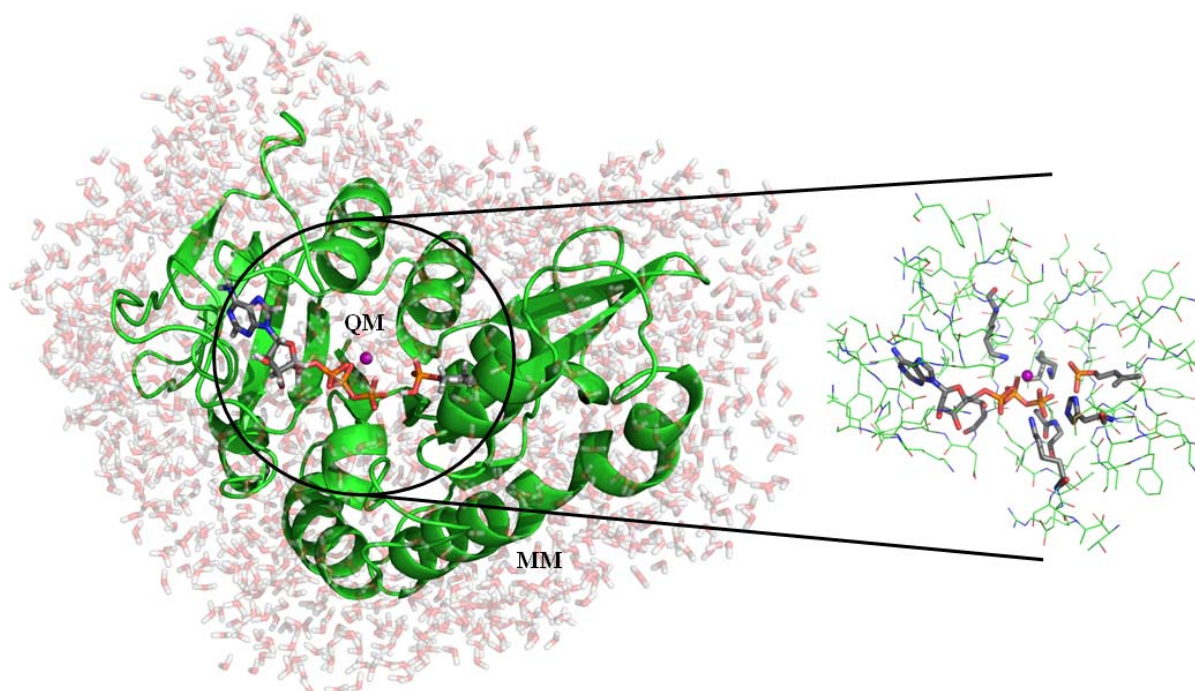
## QM/MM Optimization

Representative conformations were obtained from the MD trajectory by cluster analysis. The structure with the highest occurrence and which appeared to be most catalytically viable was optimized using a two-layer QM/MM method implemented using ONIOM in Gaussian 09.<sup>13</sup> The systems were divided into two layers, namely the high layer (QM) and the low layer (MM). The high layer was optimized at the B97D/6-31G (d) level of density functional theory while the low layer was optimized by the Amber Parm99 force field. Electronic embedding was used to integrate the point charges of MM layer into the QM region to polarize the QM wave function and a quadratically convergent Hartree-Fock (QC-SCF) method was applied.<sup>14</sup> Residues within 6 Å of the active site were free to move while the rest of the system was frozen, implemented using the Tao protocol.<sup>15</sup>

Mg<sup>2+</sup> and the substrate are included in the QM region, as well as truncated nucleotide ATP with only the methyl-triphosphate tail kept. Three coordinated water molecules along with the structural water that is H-bonded with Thr163, Lys205 and the β-phosphate oxygen of ATP were also incorporated in the QM region. In addition, the side chains of His50, Lys5, Lys14, Lys205 and Thr163 along with Gly7 and Gly8 joined by an amine linkage were included in the high layer (Figure S1). The total charge of the QM region was -1.

The RESP charges for ATP and IP used in MD simulations were adopted in the QM/MM optimization. The QM region was optimized using B97d<sup>16</sup> as the DFT functional to take into account the effect of dispersion and the 6-31g(d) basis set while the MM region was optimized by the Amber Parm 99 force field. The optimized structure was used as a starting point for a potential energy scan along the reaction coordinate, *i.e.* between the phosphorous

atom of the  $\gamma$ -phosphate of ATP and the protonated phosphate oxygen of ATP in steps of 0.1 Å. A transition state optimization was performed on a suitable structure from the PES scan. Frequency calculations were used to validate the transition state through a unique imaginary frequency. Stationary points were connected through an intrinsic reaction coordinate (IRC) calculation.



**Figure 1.** Representative cluster model of IPK saturated in solvent. The QM region is highlighted by stick mode.

### 3. RESULTS AND DISCUSSION

#### 3.1 Conformational Changes of IPK with Substrate Binding

IPK belongs to AAK kinase family, which includes N-acetyl- L-glutamate kinase (NAGK),<sup>17</sup> the aspartokinase (AK) N-terminal domain,<sup>18</sup> glutamate-5-kinase (G5K),<sup>19</sup> carbamate kinase (CK),<sup>20</sup> uridine monophosphate kinase (UMPK)<sup>21</sup> and fosfomycin kinase (FomA). In the crystal structure of *S. wedmorensis* FomA,<sup>6</sup> the magnesium ion is coordinated with the  $\alpha/\beta/\gamma$  phosphates of ATP as well as three water molecules. The location of  $\text{Mg}^{2+}$  was not resolved in the crystal structure of *Methanothermobacter thermautotrophicus* (THA) IPK (PDB Code:



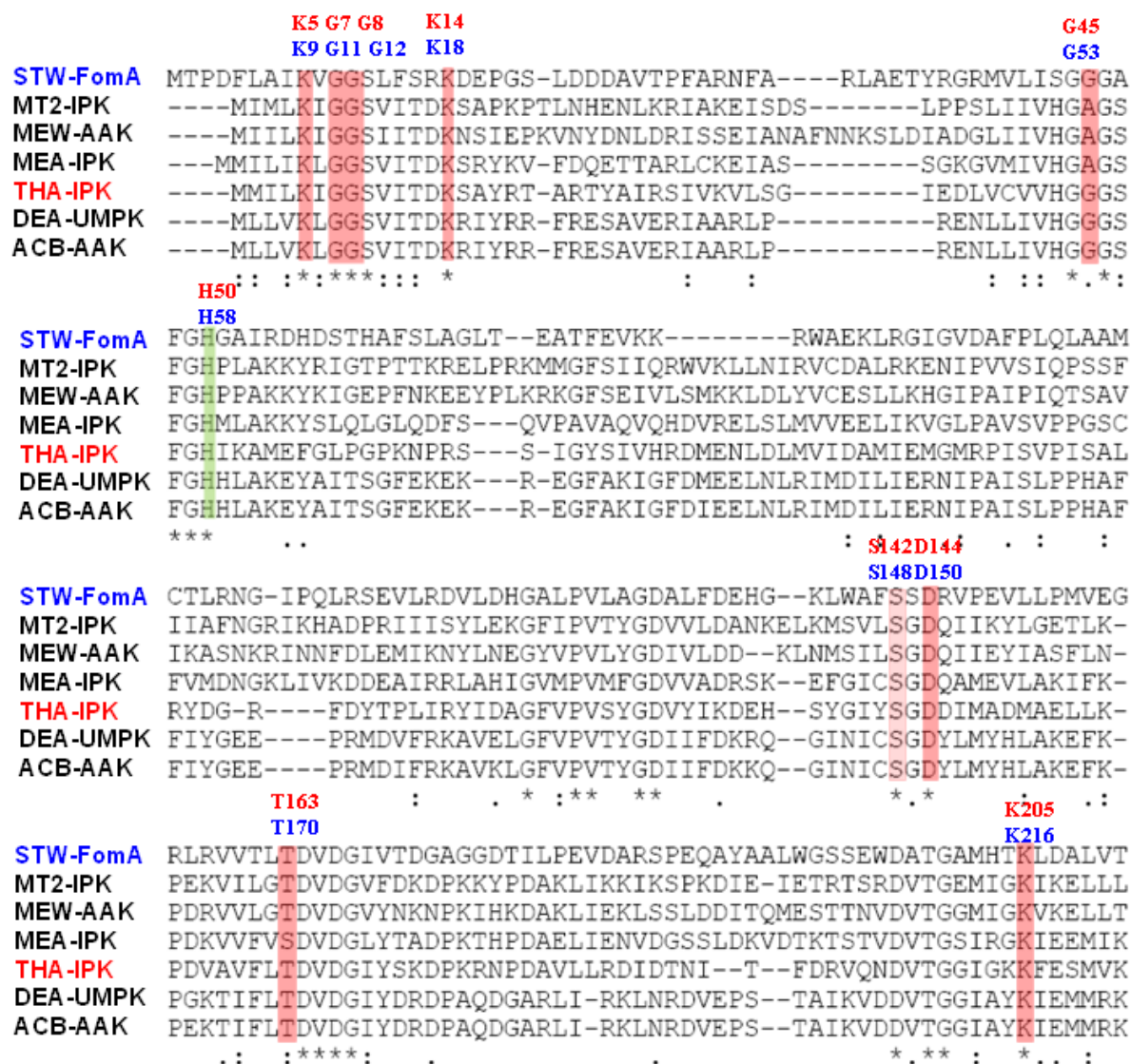
3LKK).<sup>4</sup> The  $\text{Mg}^{2+}$ -bound IPK model was built by superimposition with the FomA structure. Thus  $\text{Mg}^{2+}$  is also coordinated to the  $\alpha/\beta/\gamma$  phosphates of ATP in the IPK complex structure. The three water molecules that are coordinated to  $\text{Mg}^{2+}$  in FomA were kept to saturate the  $\text{Mg}^{2+}$  coordination in IPK.

The ambiguity in the position of atoms determined from X-ray crystallography can be resolved by molecular dynamics, which presents the dynamic conformations of the entire system with time evolution.<sup>22</sup> Here MD simulations were utilized to describe the dynamic conformation of IPK while  $\text{Mg}^{2+}$  ion is introduced in the catalytic centre. The RESP charges of the ligands in the IPK complex were calculated by single point energy calculations using HF/6-31G(d) (Supporting information Table S1). The RESP charges of ATP were collated with those reported in the Amber force field.<sup>23</sup>

RMSD analysis on the protein backbone and ligands of IPK demonstrated that the protein complex structure reached equilibrium (Figure S2). Ten representative cluster structures were obtained from cluster analysis performed for the last 10ns of the equilibrated trajectory and compared to the lowest energy structure (Table S2).

High resolution crystal structures of IPK from *Thermoplasma acidophilum* (THA) in complex with the substrate and product were resolved.<sup>4</sup> A catalytic motif ‘lysine triangle’ comprised of Lys5, Lys14 and Lys205 was identified amongst IPKs. Lys14, one of the lysine residues in a “lysine triangle”, was suggested to stabilize the transition state. Asp144 forms ionic interactions with Lys5 and Lys205. His50, a featured residue for the ‘phosphate’ subdivision of AAKs forms H- bonds with both the phosphate group of IP and the  $\gamma$ -phosphate of ATP, poising the ligands in a favourable conformation for phosphoryl transfer. The phosphate group of IP also forms a H-bond with the backbone N-H of Gly45, which is located in a well conserved glycine-rich loop  $\beta$ 2- $\alpha$ B and was suggested to stabilize the

transition states and phosphorylated products of the AAK kinases.<sup>24, 25</sup> Sequence alignment shows that these residues are highly conserved among the “phosphate” subdivision of the AAK kinases (Figure 2).



**Figure 2.** Sequence alignments of seven kinases in the “phosphate” subdivision of the AAK family using the Clustal Omega web server provided by the European Bioinformatics Institute. Key residues Lys5, Gly7, Gly8, Lys14, Asp144, Thr163 and Lys205 in THA-IPK are labelled in red; the corresponding residues in STW-FomA are labelled in blue. His50 in IPK is highlighted in green. The abbreviations of the species names and accession numbers are as follows: THA, *Thermoplasma acidophilum* (THA-IPK: CAC11251.1); STW,

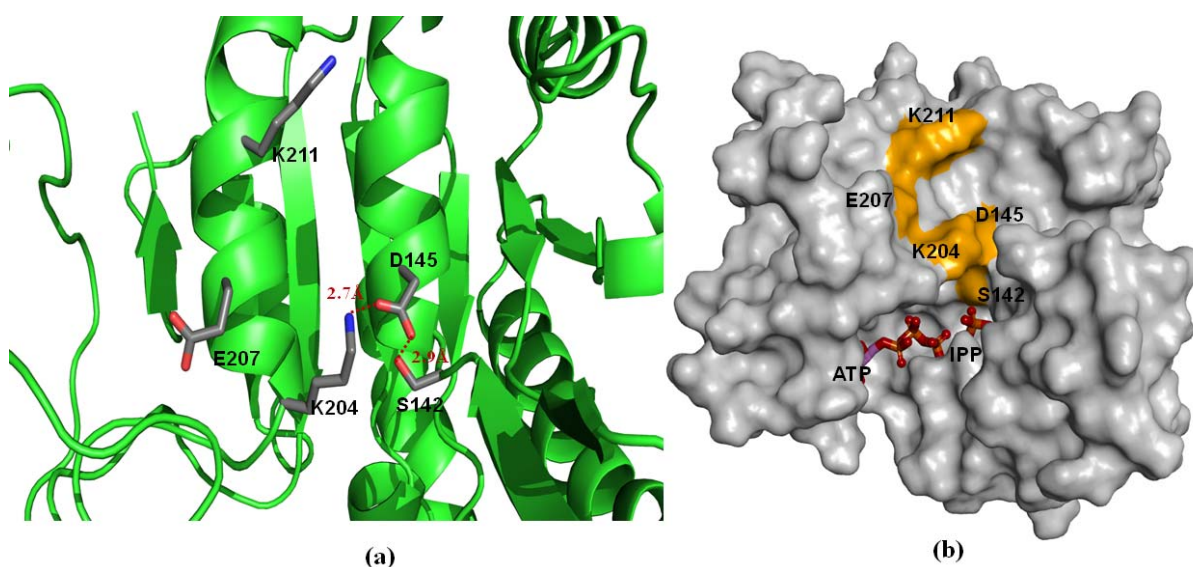
*Streptomyces wedmorensis* (STW-FomA: BAA32493.1); DEA, *Desulfurococcus amylolyticus* (DEA-UMPK: AFL65937.1). MT2, *Methanothermobacter* sp. (MT2-IPK: BAW30997.1); MEW, *Methanobrevibacter wolinii* (MEW-AAK: WP\_042707947.1); MEA, *Methanogenic archaeon* (MEA-IPK: AMK13326.1); ACB, *Aciduliprofundum boonei* (ACB-AAK: WP\_049940066.1).

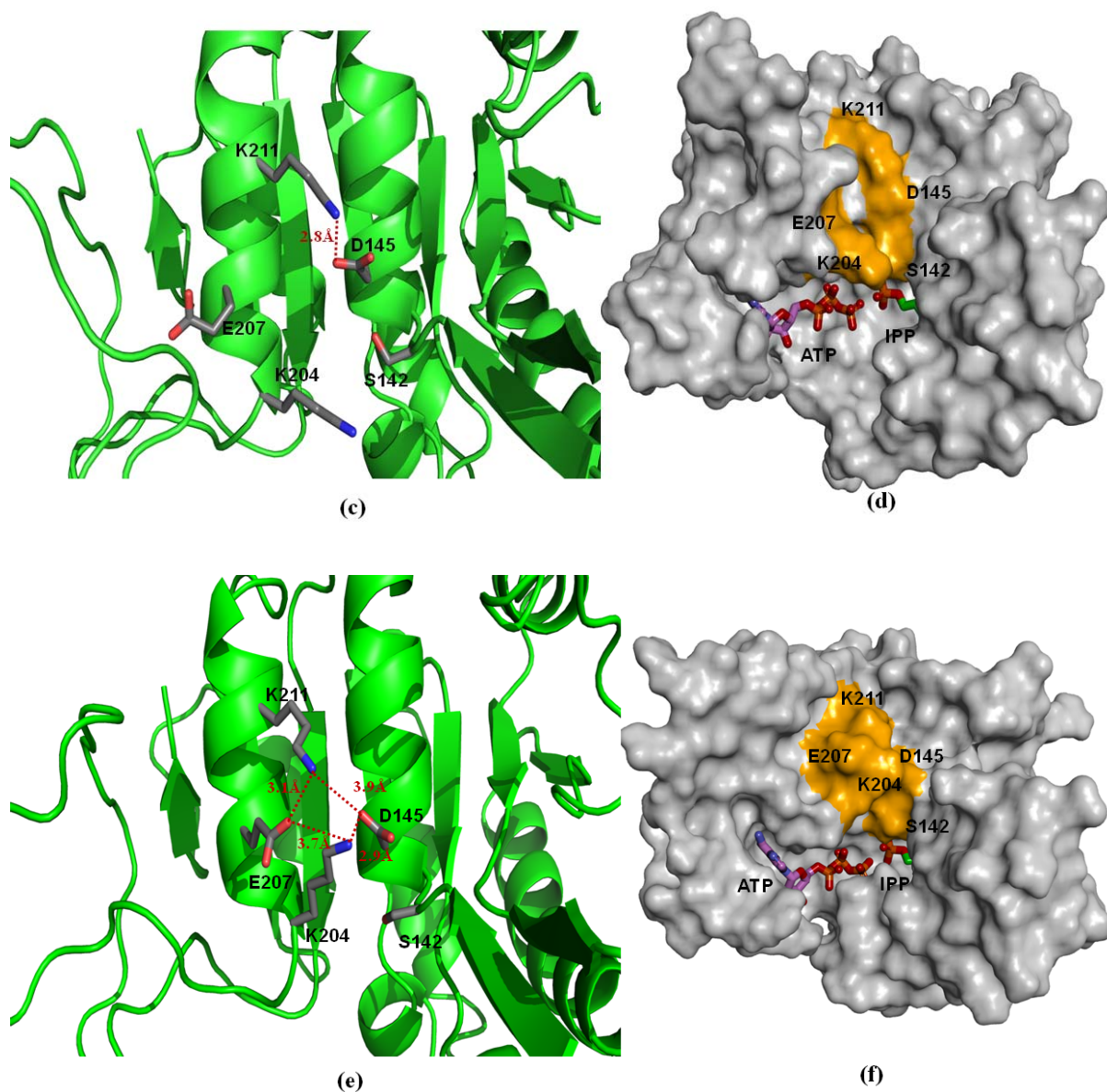
From our MD simulations, the architecture of the active site from the crystal structure was largely maintained. The  $Mg^{2+}$  remains octahedrally coordinated to the  $\alpha/\beta/\gamma$  phosphates of ATP and three water molecules. Furthermore, a coordinating water molecule forms a hydrogen bond with both Asp144 and with the  $\alpha$ -phosphate of ATP. Although no coordinating water is present in the substrate-bound IPK structures, an analogous water molecule can be identified that is H-bonded to Asp144 and  $\alpha$ -phosphate of ATP in the product-bound IPK structure (PDB code: 3LL5), <sup>4</sup> indicating the introduced coordinating water in the complex and resulting  $Mg^{2+}$  coordination environment is reasonable.

Thr163 is invariant among the “phosphate” subdivision of the AAK kinase family. In the catalytic centre of *S. wedmorensis* FomA, <sup>6</sup> the invariant residue Thr170 stabilizes ATP via the H-bond with the  $\beta$ -phosphate oxygen. However, in the crystal structure of IPK-IP-ATP (PDB: 3LKK), <sup>4</sup> Thr163 points away from the catalytic site, which represents an inactive configuration. Interestingly, with time evolution, we found Thr163 turns towards of ATP to form an H-bond with the  $\beta$ -phosphate and a structural water, further stabilizing ATP so as to facilitate the phosphorylation (Figure S3a). It is worth noting that in the crystal structure of the product-bound IPK structure, IPK-IPP-ADP (PDB: 3LL5), <sup>4</sup> Thr163 is arrayed similarly to form a H-bond with the  $\beta$ -phosphate of ATP. Thus Thr163 in IPK also plays an indispensable role to stabilize the flexible tail of ATP, and therefore facilitating the phosphoryl transfer.

The “phosphate” subdivision of AAK kinases is characterized by a conserved serine residue Ser142 on the  $\alpha$ E helix, which was suggested to stabilize the substrates.<sup>4</sup> The invariant serine residue Ser148 in FomA is H-bonded to the phosphate group of the substrate fosfomycin.<sup>6</sup> However, this stabilization function does not exist in the crystal structures nor is observed in the MD simulated structures of IPK. Instead, we found that Ser142 makes an H-bond with the carboxylate group of Asp145 (Figure S3b and Figure S5b, replica 1 & 3, Figure 3a).

It should be noted that in the crystal structures of both IP-ATP-IPK and IPP-ADP-IPK,<sup>4</sup> Lys204, Glu207 and Lys211 located on the  $\alpha$ G helix are exposed (Figure S5). During the time evolution of the MD simulations, an exposed lysine residue Lys204, which is adjacent to one of the catalytic triad residue Lys204 and close to the highly flexible exposed segment missing in the crystal structure, was found to switch inwards to form ionic bond with Asp145 (Figure S3c). This structural rearrangement further causes another exposed residue Glu207 to turn inwards, forming ionic interactions with Lys204 (Figure S3d, replica 2 & 4 and Figure 3e). In the mean time, Lys211, which is also exposed in the crystal structure, flips inwards to interact with either Asp145 (Figure S3f, replicas 1, 3 & 4, Figure 3c & e) or Glu207 (Figure S3e, replica 2 and Figure 3e), thus resulting in a strong salt bridge network (Figure 3e).



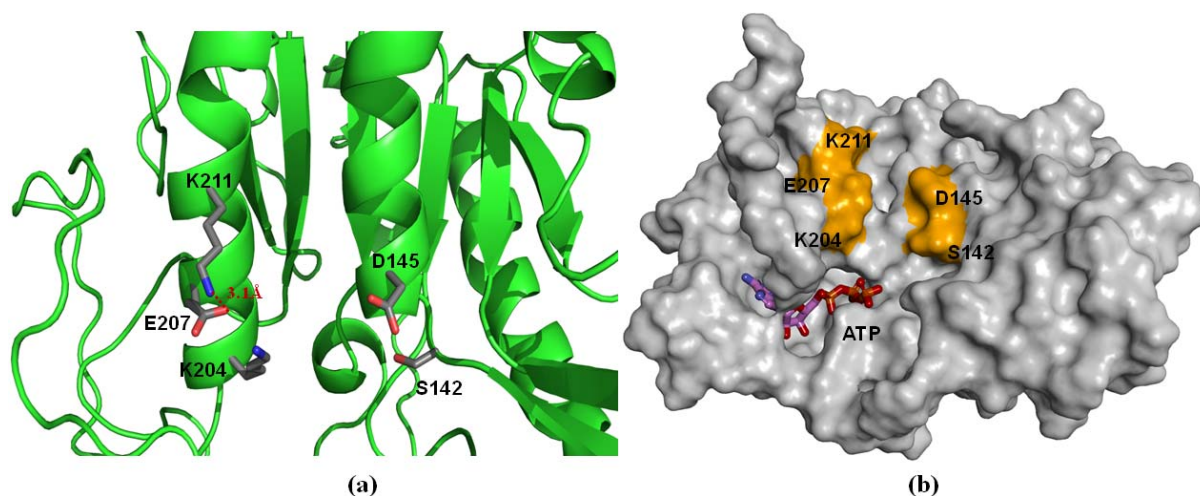


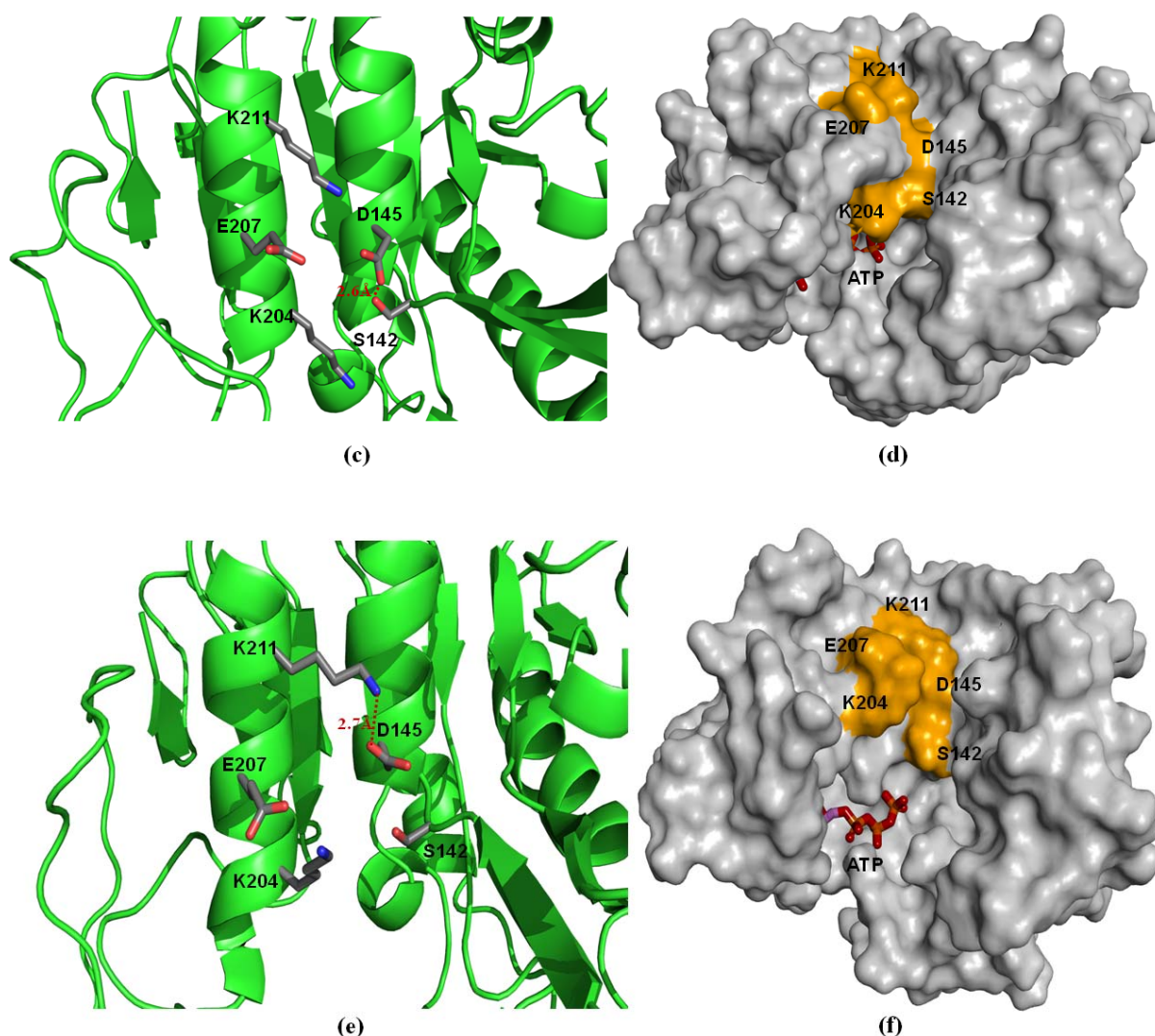
**Figure 3.** MD simulated structures of the IP-ATP-Mg<sup>2+</sup>-IPK complex (a) salt bridge formed between Lys211 and Asp145 and H-bond between Asp145 and Ser142; (c) salt bridge formed between Lys211 and Asp145; (e) salt bridge network formed among Lys211, Glu207, Lys204 and Asp145; (b), (d) & (f) are the corresponding surface of (a), (c) and (e).

In order to understand the conformational change caused by substrate binding, we also conducted MD simulations for ATP-Mg<sup>2+</sup>-IPK in absence of the substrate IP.



Thr163 was found to form H-bond with the  $\beta$ -phosphate of ATP in three of the replicas (Figure S4a, replicas 2-4), indicating the stabilization of ATP phosphate tail by Thr163 is further facilitated by the substrate binding. Similar to the substrate-bound complex, Lys211 on the  $\alpha$ G helix also approach Glu207 to form ionic interactions (Figure S4e, Figure 4a). However, the interaction between Lys204 and Glu207 observed in the substrate-bound complex is missing (Figure S4d). In addition, the ionic network formed between Asp145 on the  $\alpha$ E helix with Lys204 observed in the substrate-bound complex is lost in majority of the snapshots of the ATP-Mg<sup>2+</sup>-IPK complex (Figure S4c, replicas 1-3), which results in an open channel for the substrate to enter the active site (Figure 4b). Similar to the substrate-bound complex, Asp145 can either interact with Ser142 via an H-bond (Figure S4b, replica 4, Figure 4c) or form an ionic bond with Lys211 (Figure S4f, Figure 4e). Therefore, it appears that Ser142 and Asp145 on the  $\alpha$ E helix, together with Lys204, Glu207 and Lys211 on the  $\alpha$ G helix, may constitute a channel to mediate the substrate entry (Figure 4b). After the substrate is bound (Figure 3f), the channel is shut, locking the substrate into the catalytic site to complete the phosphoryl transfer.





**Figure 4.** MD simulated structures of the ATP-Mg<sup>2+</sup>-IPK complex (a) salt bridge formed between Glu207 and Lys211 and Lys204 on the  $\alpha$ G helix, respectively; (c) H-bond formed between Ser142 and Asp145 on the  $\alpha$ E helix; (e) salt bridge formed among Lys211 and Asp145; (b), (d) & (f) are the corresponding surface of (a), (c) and (e).

### 3.2 The Reaction Mechanism of IPK Revealed by QM/MM Calculations

Combined QM/MM studies are a powerful way to elucidating the mechanism of enzyme-catalyzed reactions.<sup>26</sup> Cluster analysis was performed for the last 10 ns of the equilibrated MD trajectories and the most catalytically favorable structure with the highest occurrence was selected for further optimization using QM/MM optimizations (Table S2). The QM

region consisted of the substrate IP,  $Mg^{2+}$  and the three water molecules saturated the coordination sphere of  $Mg^{2+}$ . ATP was truncated with only the methyl-triphosphate tail kept. The truncated catalytic triad residues Ly5, Lys14 and Lys205 as well as the two residues participating organizing the active residues His50 and Asp144 were included in the QM region. Our MD simulations disclosed that Thr163 functions to stabilize the phosphate tail of ATP and therefore it was also included in the QM region. In addition, Gly8 located at the other glycine-rich loop  $\beta 1-\eta 1$  forms a hydrogen bond with the  $\gamma$ -phosphate oxygen of ATP and another glycine residue Gly7 forms a weak C-H-O H-bond with the  $\beta$ -phosphate. Since the  $\beta$ -, $\gamma$ - phosphate groups of ATP are to be cleaved during phosphoryl transfer, these two glycine residues are also included in the high layer region. Except for the three coordinated water molecules, a fourth water molecule that formed hydrogen bonds with the  $\beta$ -phosphate oxygen of ATP.

In the optimized reactant structure of IPK, the distance between the  $\gamma$ -phosphate phosphorous atom of ATP and the phosphate oxygen of IP are 3.18 Å apart. The octahedral coordination sphere of  $Mg^{2+}$  was maintained in the QM/MM optimized reactant complex indicating the stable geometry of the catalytic site. In the optimized structure of the transition state, the phosphoryl group of ATP was 2.00 Å from the phosphate of IP. It is worth noting that the proton of the phosphate hydroxyl group of the substrate IP has transferred onto the  $\gamma$ -phosphate of ATP in the transition state. Throughout the reaction process,  $Mg^{2+}$  remains coordinated to the  $\alpha/\beta/\gamma$  phosphates of ATP. His50 retains hydrogen bonded with the substrate phosphate oxygen, aligning the substrate for phosphoryl transfer. Thr163 forms a hydrogen bond with the  $\beta$ -phosphate of ATP, which assists to position the nucleotide, facilitating the transfer of the phosphate group to the substrate.

As shown in the crystal structures of IPK, Lys5 makes strong ionic interactions with the  $\beta$ - /  $\gamma$ -phosphate of ATP and the phosphate group of the substrate, thereby aligning the substrate



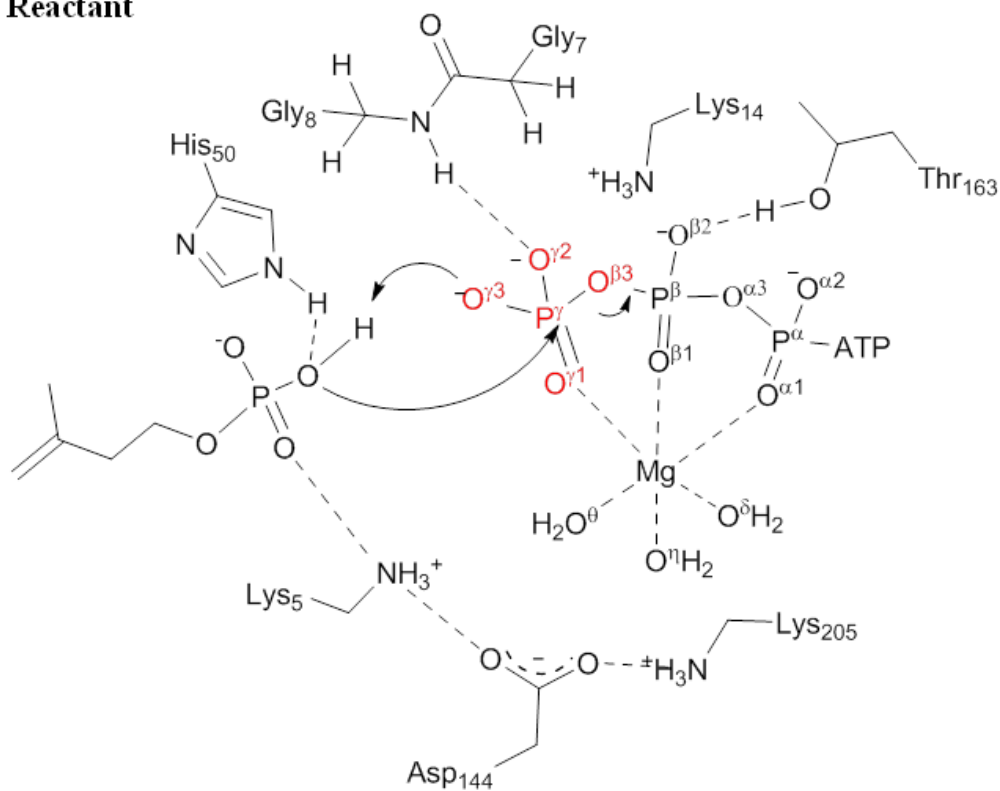
and cofactor perfectly for phosphoryl transfer to occur. This stable configuration was maintained in the QM/MM optimized structures. In addition, Lys205 maintains salt bridge interactions with the  $\beta$ -phosphate of ATP and Asp144.

Lys14 is also invariant in other AAK kinases such as FomA (Figure 2). In the crystal structure of substrate- and product- bound IPK, Lys14 is around 3.09 Å away from the  $\beta$ , $\gamma$ -bridging oxygen of ATP. The short distance indicates it may stabilize the negative charge developed around the  $\beta$ , $\gamma$ -bridging oxygen during phosphorylation. In the QM/MM optimized IPK-IP-ATP reactant complex structure, it is 3.97 Å away from the  $\beta$ , $\gamma$ -bridging oxygen and forms stable H-bonds with the  $\alpha$ -, and  $\gamma$ -phosphate instead. Interestingly, it approaches  $\beta$ , $\gamma$ -bridging oxygen in the transition state (3.61 Å) and eventually forms H-bond in the IPK-IPP-ADP product complex (2.87 Å). The short distance in the QM/MM optimized transition state provides direct evidence that Lys14 stabilizes the transition state of the phosphorylation reaction. The invariant lysine residues in other AAK kinases such as Lys18 in *Streptomyces wedmorensis* FomA <sup>6</sup> and Lys10 in *Sulfolobus solfataricus* UMPK <sup>27</sup> may play same role to stabilize their respective transition states.

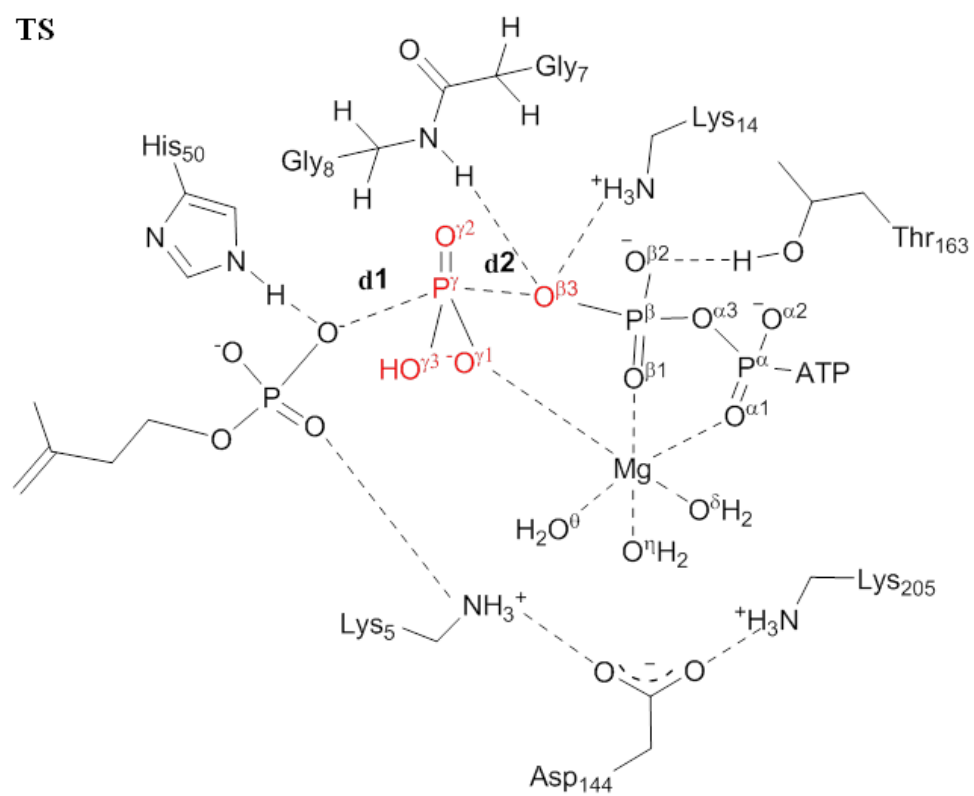
In the crystal structures of IPK, the NH group of Gly8 on the glycine-rich loop is H-bonded with the  $\gamma$ -phosphate oxygen (around 2.83 Å). Notably, it moves away from the terminal  $\gamma$ -phosphate and gradually approaches the  $\beta$ , $\gamma$ -bridging oxygen during phosphorylation, being 3.42 Å, 3.13 Å and 3.04 Å away from the bridging oxygen in the QM/MM optimized reactant, TS and product complexes, respectively. Thus we propose that Gly8 also plays a crucial role in stabilizing the transition state during phosphoryl transfer. Ser142 is H-bonded with Asp145, which in turn interacts with Lys204, Glu207 and Lys211 via an ionic interaction network. Thus Ser142 may regulate the substrate entry rather than substrate binding.

The reported  $K_{\text{cat}}$  of THA IPK for IP is  $8.0 \text{ s}^{-1}$ ,<sup>28</sup> corresponding to an activation energy of 16.89 kcal/mol. Based on our QM/MM calculations, an activation energy of 17.58 kcal/mol was estimated, which is in good agreement with the kinetic data. A benchmark study was conducted using larger basis sets for the DFT functional to evaluate the effect of diffusion and polarization. A barrier of 19.34 kcal/mol was obtained based on full geometry optimization with 6-311++G(d,p) basis set (Table S3). The little difference in the calculated barriers using the small and large basis sets indicates the 6-31G(d) basis set is sufficient for evaluating the reaction energetics. The key distances for phosphoryl transfer in the transition state are,  $d1=2.02 \text{ \AA}$  and  $d2=2.00 \text{ \AA}$  (Figure 8). The transition state was validated by a unique imaginary frequency of -158.66 corresponding to phosphoryl transfer. Furthermore, we performed an IRC calculation to authenticate the transition state and demonstrated that the TS connects the reactant and product (Figure S4).

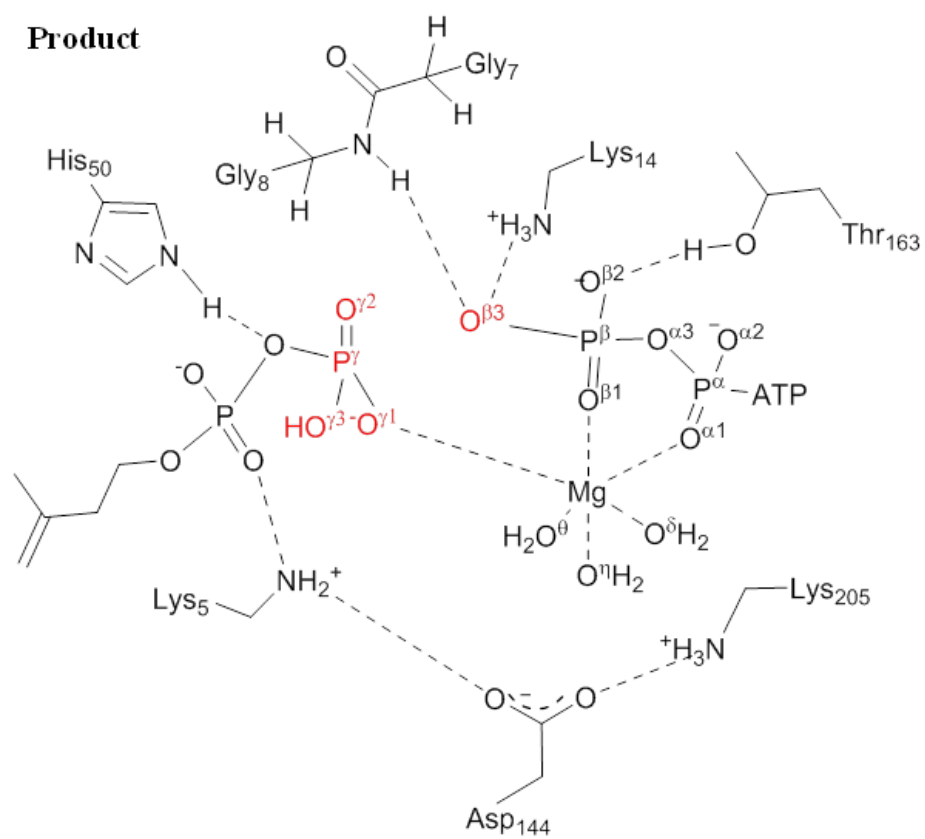
### Reactant



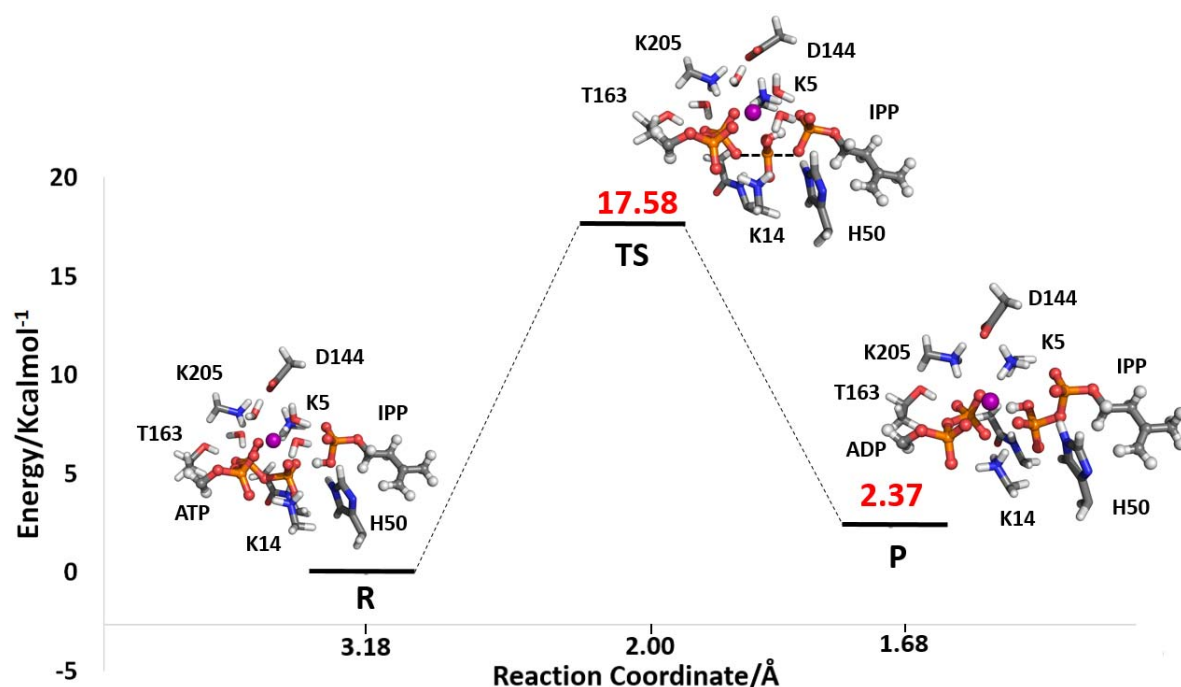
**TS**



**Product**



**Figure 5.** 2-D sketch of the phosphorylation reaction. The reaction coordinates are d1 and d2, which are corresponding to the distance between the phosphate oxygen of IP and P $\gamma$  of ATP, and the distance between the P $\gamma$  of ATP and O $\beta^3$ .

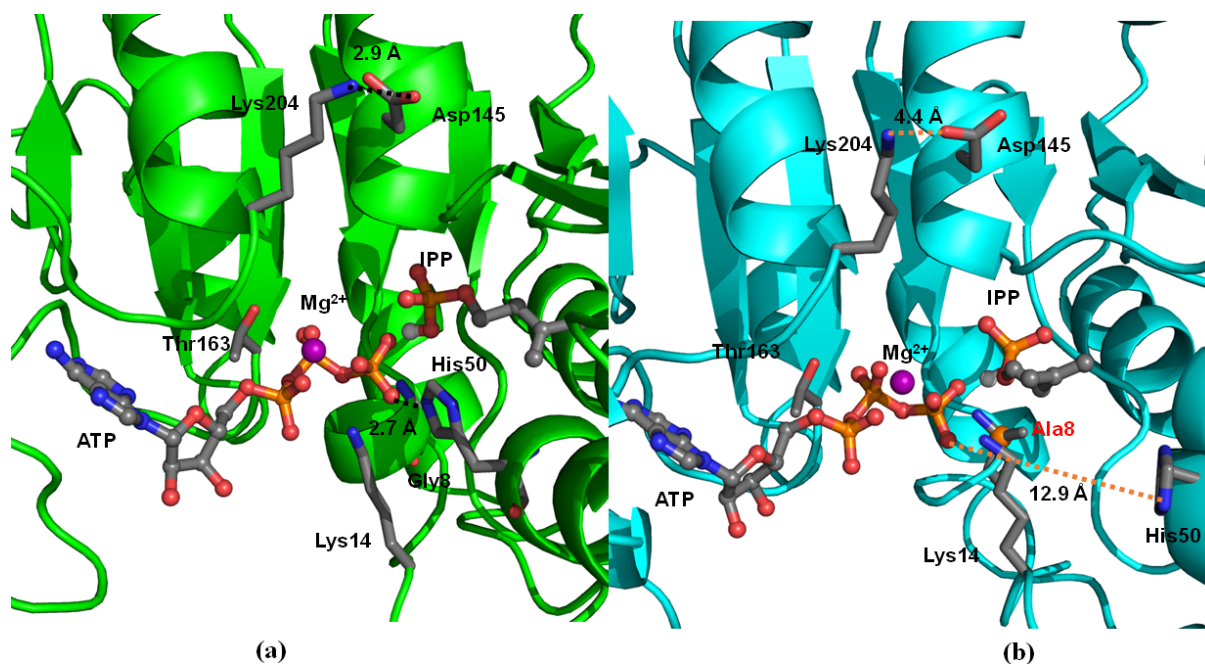


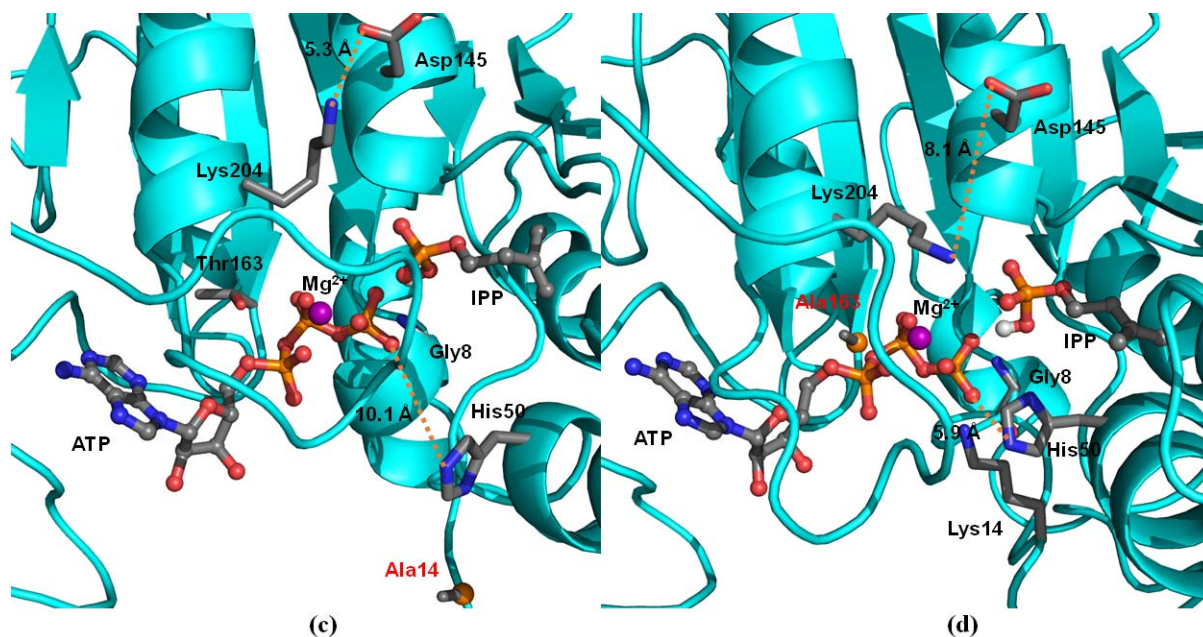
**Figure 6.** QM/MM energy profile for phosphoryl transfer from ATP to isopentenyl phosphate. Energies are given relative to the reactants (R).

Pauling's formula, <sup>29</sup>  $D(n) = D(l) - 0.6 \log n$  was used to determine the associative or dissociative nature of the transition state. Here  $D(l)$  is 1.73, the average length of a P-O bond and  $D(n)$  is the average of  $r_{O-P\gamma}$  and  $r_{P\gamma-O\beta^3}$  in the transition state structure. The estimated fractional bond number ( $n$ ) was 0.34, which indicates a dissociative character of 66%. The similarity of the over-all fold and the substrate and nucleotide binding environments (Figure 2) between IPK and other AAK kinases in the “phosphate” subdivision such as FomA suggests they may share similar mechanisms.

### 3.3 Impact of Mutation of Catalytically Relevant Residues in IPK

Among all the AAK family members, FomA is most structurally similar to IPK. <sup>6</sup> Kinetic studies revealed that a K18A mutant in FomA (corresponding to the conserved residue K14 in IPK) resulted in no detectable activity in the enzyme. In order to validate that Lys14 is also crucial for the substrate binding in IPK and the kinase's activity, MD simulation was run for the K14A variant. From our MD simulations of the K14A mutant it can be seen that the hydrogen bonds formed between His50 and the phosphate group of IP and the  $\gamma$  phosphate of ATP in the WT kinase are now broken (Figure 7c and Figure S7). As a result, His50 is pulled away from the active site of IPK and forms a hydrogen bond with the side chain of Asn196. Thus it is evident that Lys14 plays a crucial role in maintaining a catalytically competent active site architecture.





**Figure 7.** Representative structure of the ATP-Mg<sup>2+</sup>-IPK complex from the cluster analysis of the last 10ns of 100-ns MD trajectory (a) WT (b) G8A (c) K14A (d) T163A.

Based on our QM/MM study on the catalytic mechanism of IPK Gly8 was suggested to play an important role in stabilizing the TS of the reaction mechanism. We also run MD simulation for G8A and found a notable conformational change for His50, which is shifted away from the active site. The steric hindrance generated by mutating glycine into a bulkier alanine group pushed His50 away from the active site. Further, the H bond formed between the NH group of Gly8 and the  $\beta$ , $\gamma$ -bridging oxygen of ATP in the WT kinase is lost when Gly8 is replaced by alanine (Figure 7b and Figure S7). This indicates that the flexibility induced by Gly8 in the WT kinase is essential to tether the catalytic His50 in the proximity of the catalytic pocket, enabling the phosphorylation reaction to occur.

From the MD simulations of the WT ATP-Mg<sup>2+</sup>-IPK complex we found that Thr163 plays a crucial role in stabilizing the flexible tail of ATP by forming a H bond with the  $\beta$ -phosphate of the nucleotide. To validate this effect MD simulation was performed for a T163A mutant.

We found that similarly to the K14A and G8A mutants, catalytic residue His50 no longer maintains its position in the active site of IPK (Figure 7d and Figure S7). Since the hydroxyl group of threonine side chain is replaced with a methyl group of alanine no hydrogen bond is formed with the  $\beta$ - phosphate of ATP. Interestingly, we found the ionic bridge formed by Asp145 and Lys204 observed in the WT kinase is lost in all these mutants. This further indicates the non-conserved second sphere residues such as Asp145 and Lys204 play indispensable role in stabilizing the substrate in the catalytic site.

To evaluate the role of the second sphere residues K211, S142 and K211 in the substrate binding, we also run MD simulations for the K211A, S142A and K211A variants. We found that His50 moves away in K211A; In S142A and K211A, the salt bridge between Asp145 and Lys204 is broken, although His50 remains in proximity to the ATP and IP substrates (Figure S8).

## CONCLUSIONS

The biosynthesis of isopentenols provides a promising approach for the production of environmentally benign biofuels. IPK catalyses the  $Mg^{2+}$ -ATP dependent phosphorylation of IP to produce IPP, an important precursors of isopentenols. However, the electron density of a  $Mg^{2+}$  ion could not be unambiguously assigned to crystal structure of IPK, which has limited our understanding the catalytic environment and the mechanism of the phosphorylation reaction.

In the present study, we built a  $Mg^{2+}$ -bound IPK-IP-ATP complex model based on the location of  $Mg^{2+}$  in another AAK kinase, FomA, and elucidated that the phosphorylation reaction mechanism of IPK using combined MD simulations and hybrid QM/MM calculations. His50 is a characteristic residue in the “phosphate” subdivision of AAK kinase family which holds the phosphate groups of both ATP and IP substrates in the catalytic site,

facilitating the reaction to happen. Thr163 also plays an indispensable role to stabilize the flexible tail of ATP, and therefore facilitating the phosphoryl transfer.

Lys204, together with Glu207 and Lys211 on the  $\alpha$ E helix, undergo significant conformational change during substrate binding, forming a strong ionic network with Asp145, which in turn tethers the invariant Ser142 via H-bond interaction. Substrate binding induces shutting of the substrate entrance channel formed between the  $\alpha$ E and  $\alpha$ G helices.

Furthermore, we propose that a highly conserved residue Gly8 on a glycine-rich loop, together with Lys14 on the catalytic lysine triangle, plays a crucial role in stabilizing the transition state during phosphoryl transfer. These findings provide the atomic-level basis for rational bioengineering of IPK or other related kinase for the production of biofuels.

## **AUTHOR INFORMATION**

### **Corresponding Author**

\*E-mail: [m.huang@qub.ac.uk](mailto:m.huang@qub.ac.uk)

## **ACKNOWLEDGEMENTS**

J.M. acknowledges the financial support from the Department of Education and Learning (DEL), Northern Ireland. The authors acknowledge the financial support from INVEST NI Research and Development Programme, part financed by the European Regional Development Fund under the Investment for Growth and Jobs programme 2014-2020. We are grateful for the computing resources from QUB high performance computing centre.

## **ASSOCIATED CONTENT**

### **Supporting Information**

The supporting information is available free of charge on the ACS Publications website.



RESP charges of the ligands (Table S1); Cluster analysis of MD trajectory of the IP-ATP-IPK-Mg<sup>2+</sup> complex (Table S2); benchmark study on the basis set for the DFT functional in the QM region of the QM/MM study (Table S3); QM region of the QM/MM calculations (Figure S1); Plots of important distances throughout MD simulations (Figure S2-S4); H-bond interaction and surface representation of IP-ATP-IPK complex (Figure S5); IRC graph (Figure S6); comparison of important distance changes in the WT and variants (Figure S7); Representative snapshots of IPK variants (Figure S8).

## REFERENCES

- (1) Zheng, Y.; Liu, Q.; Li, L.; Qin, W.; Yang, J.; Zhang, H.; Jiang, X.; Cheng, T.; Liu, W.; Xu, X. Metabolic Engineering of Escherichia Coli for High-Specificity Production of Isoprenol and Prenol as next Generation of Biofuels. *Biotechnol. Biofuels* **2013**, *6*, 57.
- (2) Levy, S. B.; Marshall, B. Antibacterial Resistance Worldwide: Causes, Challenges and Responses. *Nat. Med.* **2004**, *10* (12 Suppl), S122-129.
- (3) Dellas, N.; Noel, J. P. Mutation of Archaeal Isopentenyl Phosphate Kinase Highlights Mechanism and Guides Phosphorylation of Additional Isoprenoid Monophosphates. *ACS Chem. Biol.* **2010**, *5*, 589–601.
- (4) Mabanglo, M. F.; Schubert, H. L.; Chen, M.; Hill, C. P.; Poulter, C. D. X-Ray Structures of Isopentenyl Phosphate Kinase. *ACS Chem. Biol.* **2010**, *5*, 517–527.
- (5) Huang, M.; Wei, K.; Li, X.; McClory, J.; Hu, G.; Zou, J.-W.; Timson, D. Phosphorylation Mechanism of Phosphomevalonate Kinase: Implications for Rational Engineering of Isoprenoid Biosynthetic Pathway Enzymes. *J. Phys. Chem. B* **2016**, *120*, 10714–10722.

- (6) Pakhomova, S.; Bartlett, S. G.; Doerner, P. A.; Newcomer, M. E. Structural and Biochemical Insights into the Mechanism of Fosfomycin Phosphorylation by Fosfomycin Resistance Kinase FomA. *Biochemistry* **2011**, *50*, 6909–6919.
- (7) Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (8) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of PKa Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295.
- (9) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- (10) Case, D. A.; Babin, V.; Berryman, J.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Iii, C.; E, T.; Darden, T. A.; Duke, R. E.; *et al.* Amber 14. **2014**.
- (11) Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; *et al.* Gaussian 09, revision D. 01. **2009**.
- (12) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (13) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and  $\text{Pt}(\text{P}(\text{t-Bu})_3)_2 + \text{H}_2$  Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357–19363.

- (14) Bacskay, G. B. A Quadratically Convergent Hartree—Fock (QC-SCF) Method. Application to Closed Shell Systems. *Chem. Phys.* **1981**, *61*, 385–404.
- (15) Tao, P.; Schlegel, H. B. A Toolkit to Assist ONIOM Calculations. *J. Comput. Chem.* **2010**, *31*, 2363–2369.
- (16) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (17) Ramón-Maiques, S.; Marina, A.; Gil-Ortiz, F.; Fita, I.; Rubio, V. Structure of Acetylglutamate Kinase, a Key Enzyme for Arginine Biosynthesis and a Prototype for the Amino Acid Kinase Enzyme Family, during Catalysis. *Structure* **2002**, *10*, 329–342.
- (18) Liu, X.; Pavlovsky, A. G.; Viola, R. E. The Structural Basis for Allosteric Inhibition of a Threonine-Sensitive Aspartokinase. *J. Biol. Chem.* **2008**, *283*, 16216–16225.
- (19) Marco-Marín, C.; Gil-Ortiz, F.; Pérez-Arellano, I.; Cervera, J.; Fita, I.; Rubio, V. A Novel Two-Domain Architecture within the Amino Acid Kinase Enzyme Family Revealed by the Crystal Structure of Escherichia Coli Glutamate 5-Kinase. *J. Mol. Biol.* **2007**, *367*, 1431–1446.
- (20) Marina, A.; Alzari, P. M.; Bravo, J.; Uriarte, M.; Barcelona, B.; Fita, I.; Rubio, V. Carbamate Kinase: New Structural Machinery for Making Carbamoyl Phosphate, the Common Precursor of Pyrimidines and Arginine. *Protein Sci. Publ. Protein Soc.* **1999**, *8*, 934–940.
- (21) Marco-Marín, C.; Gil-Ortiz, F.; Rubio, V. The Crystal Structure of Pyrococcus Furiosus UMP Kinase Provides Insight into Catalysis and Regulation in Microbial Pyrimidine Nucleotide Biosynthesis. *J. Mol. Biol.* **2005**, *352*, 438–454.
- (22) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. A Practical Guide to Modelling Enzyme-Catalysed Reactions. *Chem. Soc. Rev.* **2012**, *41*, 3025–3038.

- (23) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of Polyphosphate Parameters for Use with the AMBER Force Field. *J. Comput. Chem.* **2003**, *24*, 1016–1025.
- (24) Pakhomova, S.; Bartlett, S. G.; Augustus, A.; Kuzuyama, T.; Newcomer, M. E. Crystal Structure of Fosfomycin Resistance Kinase FomA from *Streptomyces Wedmorensis*. *J. Biol. Chem.* **2008**, *283*, 28518–28526.
- (25) Gil-Ortiz, F.; Ramón-Maiques, S.; Fita, I.; Rubio, V. The Course of Phosphorus in the Reaction of N-Acetyl-l-Glutamate Kinase, Determined from the Structures of Crystalline Complexes, Including a Complex with an AlF<sub>4</sub><sup>−</sup> Transition State Mimic. *J. Mol. Biol.* **2003**, *331*, 231–244.
- (26) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.
- (27) Jensen, K. S.; Johansson, E.; Jensen, K. F. Structural and Enzymatic Investigation of the *Sulfolobus Solfataricus* Uridylate Kinase Shows Competitive UTP Inhibition and the Lack of GTP Stimulation. *Biochemistry* **2007**, *46*, 2745–2757.
- (28) Mabanglo, M. F.; Pan, J.-J.; Shakya, B.; Poulter, C. D. Mutagenesis of Isopentenyl Phosphate Kinase To Enhance Geranyl Phosphate Kinase Activity. *ACS Chem. Biol.* **2012**, *7*, 1241–1246.
- (29) Mildvan, A. S. Mechanisms of Signaling and Related Enzymes. *Proteins* **1997**, *29*, 401–416.



## Table of Contents Graphic

